Identification of Candidate Markers on BTA14 Affecting Conformation and Functional Traits in Canadian Holstein

E. Marques[*], Z. Wang[#], G. Plastow[#] and S. Moore[#]

Introduction

Whole Genome Association (WGA) analysis seems to be the method of choice when dealing with high-density SNP panels such as the Bovine50K SNP-chip. This method is practical, because it can easily be applied to all of the SNPs in the panel, without the need to subdivide the SNPs by chromosome and then perform the laborious linkage analysis. Practicality aside, this method does not come without criticisms. At least in human studies, where these high density SNP-chips have been available for some time, WGA has yet to crack the black box underlying the genetic variation of important diseases, as it promised (Armstrong *et al*., 2009; Yang *et al*., 2007). The biggest criticisms of WGA stem from the lack of biological link between markers showing statistical association in WGA and the complex trait, lack of transferability of SNP association from one population to another (Frazer *et al*., 2009) and inadequate sample size (NCI-NHGRI Working Group *et al*., 2007). Rather than using WGA as a stand-alone procedure, it should be used as a complementary tool to linkage analysis to crack the genetic underpinnings of complex traits.

One way to circumvent the issue of multiple testing, often observed in these high-density analysis is to first identify markers whose effects are redundant in the population. A simple analysis of effect is to calculate the amount of linkage disequilibrium (LD) among markers and to select one marker from each group, thereby decreasing the number of SNPs analyzed without the loss of information. In addition, markers can be selected on the basis heterozygosity, ensuring that most of the sires are segregating for that particular marker. This approach was used in this study to identify markers on bovine chromosome 14 (BTA14) affecting conformational and functional traits in Canadian Holstein.

Material and methods

Animal Resource and Phenotypes. Two hundred and eighty Holstein bulls provided by Semex Canada were used in this study. Pedigree information was obtained from www.holstein.ca/english/Animlainq/animalinq.asp. Nine traits were analyzed in this study: calving ease (CE), daughter fertility (DF), mammary system (MS), milking speed (MilkS),

[*] Beefbooster Inc., Calgary, Alberta, T1Y5E4
[#] University of Alberta, Edmonton, Alberta, T6E2C8

somatic cell score (SCS), herd life (HL), feet and legs (FL), conformation and maternal calving ease.

QTL Mapping. Two-hundred and eighty bulls belonging to 7 half-sib families (range 24 to 36 progeny) were used for the across family quantitative trait loci (QTL) analysis which used the multiple marker interval mapping approach described by Knott *et al*. (1996). Conditional probabilities of allele inheritance were obtained from QTL Express (Seaton *et al*., 2002) and modeled using SAS (SAS Inst Inc., Cary, NC). Phenotypic data used estimated breeding values (EBVs) obtained from the Holstein Canada website (www.holstein.ca/english/Animlainq/animalinq.asp). Conditional probabilities were used as fixed effects, *DGAT1* genotype as a covariate and reliability as a weighing factor to account for the differences in the number of daughters of each sire contributing to his breeding value.

Statistical analyses. Allele substitution effects were estimated using Procedure Mixed in SAS (SAS Inst. Inc, Cary, NC) under the same statistical model as the QTL analysis described above by regressing the phenotypes on the number of copies of one allele for each SNP .

Compilation of Markers. Haplotypes were generated according to procedures described by Marques *et al*. (2008). HAPLOVIEW (Barret *et al*., 2005) was used to calculate LD and heterozygosity. A total of 502 SNP markers were used to generated pairwise LD information for the purpose of selecting non-redundant markers for a QTL scan. SNP selection based on the procedure describe above yielded 139 markers for QTL scan. SNP genotypes were generated by high throughput Illumina beadstation 500G genotyping system. *DGAT1* primers were designed according to Kaupe *et al*. (2004) and 3730 ABI sequencer was used to generated *DGAT1* genotypes for all animals.

Results and discussion

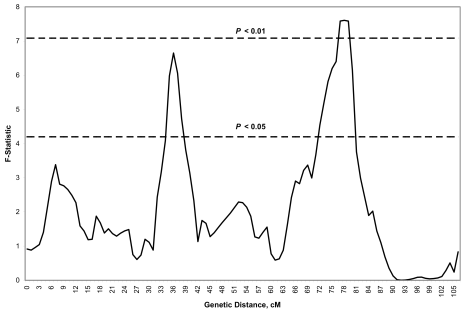
The QTL analysis performed was similar to the one described by Schnabel *et al*., (2005) which included *DGAT1* as a genetic cofactor to account for the gene's known effect on some functional traits. Out of the nine traits analyzed, 7 (CE, HL, MS, DF, FL, MilkS, SCS) yielded significant QTL peak (1% or 5% significance). Table 1 lists the QTL locations, effects and number of SNPs under the QTL. Figure 1 depicts the QTL profiles of 4 of the traits.

Table 1. Quantitative Trait Loci (QTL) on bovine chromosome 14 (BTA14) using Holstein cattle across seven families

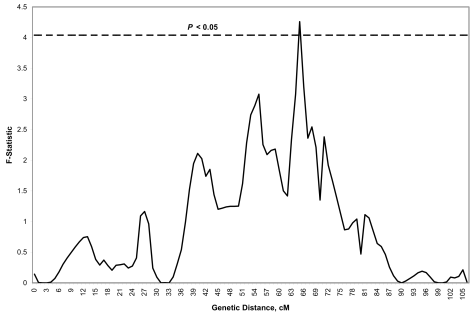
Trait	centiMorgan	Effect	SE	*P*-Value	# SNPs (P<0.01)
Calving Ease	36	1.55	0.61	0.05	5
	78	1.60	0.59	0.01	7
Daughter Fertility	65	0.92	0.45	0.05	3
Mammary System	28	1.64	0.79	0.05	1
Milking Speed	39	-1.30	0.58	0.05	0

Somatic Cell Score	37	-0.15	0.04	0.01	4
	70	-0.16	0.04	0.01	3
Herd Life	52	0.06	0.03	0.05	2
	65	0.06	0.03	0.05	7
Feet & Legs	71	-1.15	0.60	0.05	1
	85	-1.30	0.60	0.05	2

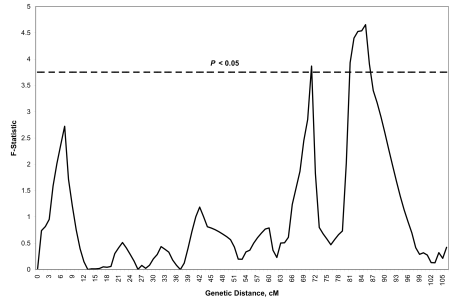
a) Calving ease



b) Daughter fertility



c) Feet & Legs



d) Somatic cell score

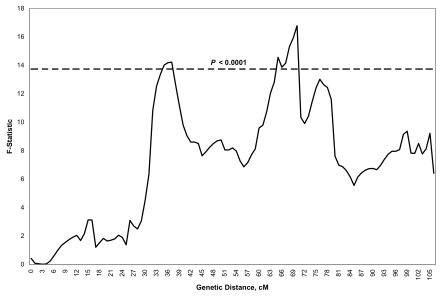


Figure 1: Across family F-statistic quantitative trait loci profiles on bovine chromosome 14 using 139 single nucleotide polymorphisms using *DGAT1* as a genetic cofactor.

In order to minimize the number of false positives, only SNPs under the QTL reaching the 1% significance threshold were evaluated. Among all the traits listed in Table 1, milking speed was the only one, which did not yield any SNPs at the 1% threshold. Calving ease and herd life both yielded QTL peaks with 7 significant ($P < 0.01$) SNPs. Among all of the 35 SNPs under the QTL, only one was common among all of the traits (Daughter Fertility and Herd Life, 67.8 cM). This SNP is located near the mitochondrial ATP synthase gene, which

catalyzes ATP synthesis by utilizing an electrochemical gradient of protons across the inner membrane during oxidative phosphorylation.

Daughter fertility and somatic cell score were the two traits with the lowest P-value markers (DF: 1.77×10^{-5}; SCS:1.48×10^{-6}). In the case of SCS, the SNP maps to a gene known for regulating defense mechanisms, confirming the biological link between the trait and SNP. In the DF case, the SNP maps to to a gene with a role in energy expenditure and higher body weight. Although the link is less obvious, it is still a plausible candidate considering the links between cow fertility and body weight (Roche *et al*., 2009).

Conclusions

Combining linkage and association analysis with linkage disequilibrium makes a stronger case for causality. WGA studies have the multiple testing problems, and unless the sample size is large, the chance of reporting false positives increases. The approach used here, though laborious confirmed its usefulness specially when using high-density SNP panels. Future analysis will include expanding this procedure to other chromosomes and to further evaluate the markers affecting conformation and functional traits in Canadian Holstein cattle.

References

Armstrong D. L., Reiff, A., Myones, B. L. et al. (2009). *Genes and Imm*., 10:446-456.

Barret, J. C., Fry, B., Maller, J. *et al*. (2005). Bioinformatics, 21:263-265.

Frazer, K. A., Murray, S. S., Schork, N. J. *et al*. (2009). *Nat Rev Genet.,* 10: 241-251.

Kaupe, B., Winter, A., Fries, R. *et al*. (2004). *J. Dairy Res*, 71:182-187.

Knott, S. A., Elsen, J. M., and Haley, C.S. (1996). *Theoretical and Applied Genetics*, 93: 71-80.

Marques, E., Schnabel, R.D., Stothard, P. *et al*. (2008). *BMC Genet*, 9: 45.

NCI-NHGRI Working Group on Replication in Association Studies, Chanock, S. J., Manolio, T. et al. (2007). *Nature*, 447: 655-60.

Roche, J. R., Friggens, N. C., Kay, J. K. *et al*. (2009). *J. Dairy Sci*., 92:5769-5801.

Schnabel, R.D, Sonstegard, T. S., Taylor, J. F. *et al*. (2005). *Anim Genet*., 36: 408-416.

Seaton, G., Haley, C. S., and Knott, S. A. *et al*. (2002). *Bioinformatics,* 18: 339-340.

Yang, Q., Biernacka, J. M., Chen, M. H. *et al*. (2007). *Genetic Epidemiology*, 31: S34-42.